Supplementary information for:

# Conserved residues control the T1R3-specific allosteric signaling pathway of the mammalian sweet taste receptor

Jean-Baptiste Chéron[1], Amanda Soohoo[2,3], Yi Wang[2,4], Jérôme Golebiowski[1], Serge Antonczak[1], Peihua Jiang[2*], Sébastien Fiorucci[1*]

[1] Université Côte d'Azur, CNRS, Institut de Chimie de Nice UMR7272, 06108 Nice, France

[2] Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA, 19104, United States

[3] Massachusetts General Hospital, Harvard Medical School, Boston, MA, 02114, United States

[4] Department of Ecology and Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan, 430072, China

*Corresponding authors: Sebastien.Fiorucci@unice.fr, pjiang@monell.org
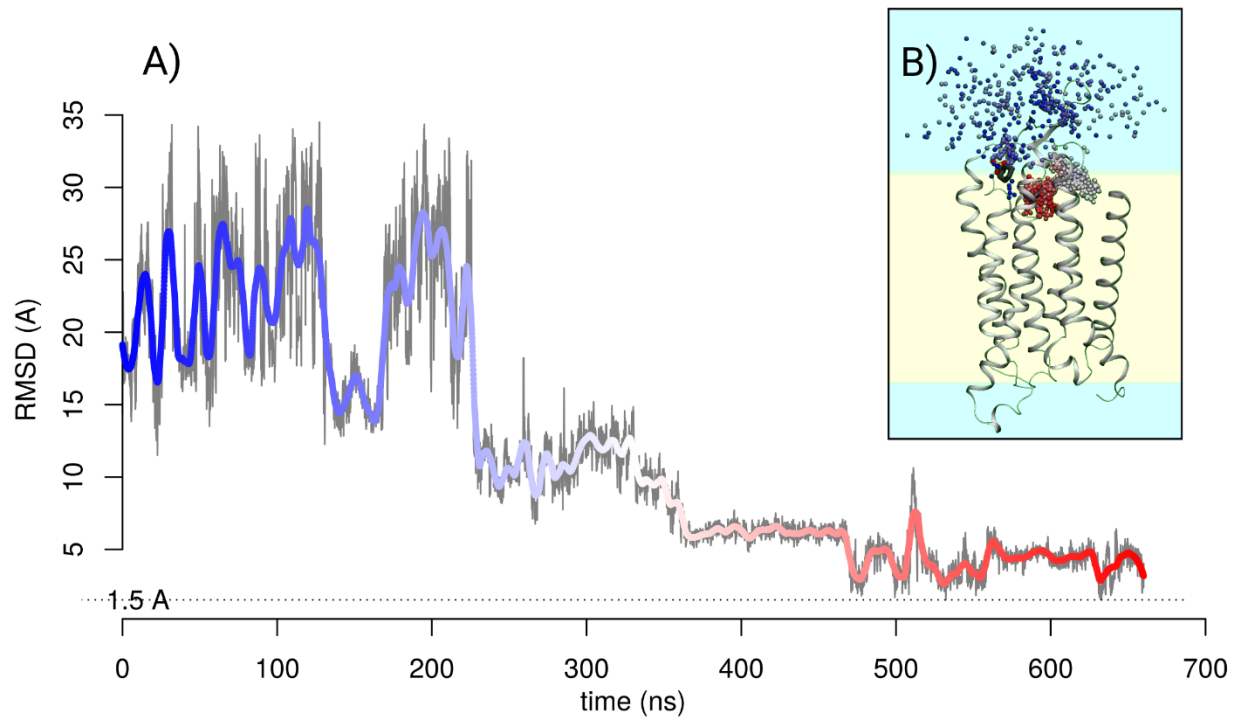
## Content:

Figure s1: Trajectory analysis of the cyclamate entering the T1R3 transmembrane domain.

A) Root mean square deviation of the cyclamate molecule along the molecular dynamics simulations with respect to the docked position. B) View of the ligand entrance pathway from the bulk phase (in blue) to the binding cavity (in red).

## Table s1: Structure-function relationships of residues involved in the cyclamate-specific signaling pathway.

Sequence conservation (less conserved residues in bold) has been calculated according to multiple sequence alignment of 41 mammalian organisms (Figure s8). Residue location is mentioned considering that the upper part of the binding site is oriented toward the extracellular medium.

| Residue | Ballesteros Weinstein numbering | Conservation | Location in the allosteric binding site | Function |
|---------|--------------------------------|--------------|------------------------------------------|----------|
| F624 | 2x56 | 97% | Upper part | cyclamate binding |
| Q636 | 3x32 | 97% | Upper part | cyclamate binding |
| Q637 | 3x33 | 94% | Upper part | cyclamate binding |
| **S640** | **3x36** | **38%** | **Upper part** | **cyclamate binding** |
| H641 | 3x37 | 100% | Upper part | cyclamate binding |
| L644 | 3x40 | 97% | Lower part | - |
| T645 | 3x41 | 97% | Lower part | - |
| L648 | 3x44 | 100% | Lower part | - |
| H721 | ecl2 | 94% | Upper part | cyclamate binding |
| **R723** | **ecl2** | **68%** | **Upper part** | **cyclamate binding** |
| **T724** | **ecl2** | **35%** | **Upper part** | **cyclamate binding** |
| **R725** | **ecl2** | **59%** | **Upper part** | **cyclamate binding** |
| S726 | ecl2 | 97% | Upper part | cyclamate binding |
| S729 | 5x39 | 88% | Upper part | cyclamate binding |
| **F730** | **5x40** | **79%** | **Upper part** | **cyclamate binding** |
| H734 | 5x44 | 100% | Lower part | - |
| N737 | 5x47 | 100% | Lower part | molecular switch |
| Y771 | 6x46 | 97% | Lower part | mlecular switch |
| W775 | 6x50 | 100% | Lower part | molecular switch |
| F778 | 6x53 | 94% | Upper part | cyclamate binding |
| **L782** | **6x57** | **62%** | **Upper part** | **cyclamate binding** |
| **R790** | **7x28** | **38%** | **Upper part** | **cyclamate binding** |
| Q794 | 7x32 | 97% | Upper part | cyclamate binding |
| C801 | 7x39 | 100% | Lower part | - |

## Table s2: Summary of molecular dynamics simulations.

Molecular Dynamics (MD) simulation lengths and average Root Mean Square Deviations (RMSD) of wild-type (wt) and mutant (mut) protein in the presence (ago) or not (apo) of cyclamate. Two independent simulations (s1 and s2) have been performed for the apo and ago systems. The RMSD is computed using the backbone atoms of the protein and the heavy atoms of the cyclamate molecule after a structural superimposition of the receptor backbone atoms to the first frame of the MD trajectory.

| | | Simulation Time (μs) | Protein RMSD (Å) | Ligand RMSD (Å) |
|---|---|---|---|---|
| **wt-ago** | **s1** | 4.12 | 2.0 | 3.4 |
| | **s2** | 3.76 | 2.9 | 3.8 |
| **wt-apo** | **s1** | 4.01 | 1.9 | - |
| | **s2** | 2.02 | 1.9 | - |
| **mut-ago** | **R725A** | 1.52 | 1.7 | 6 |
| | **W775H** | 1.00 | 1.9 | 4.7 |
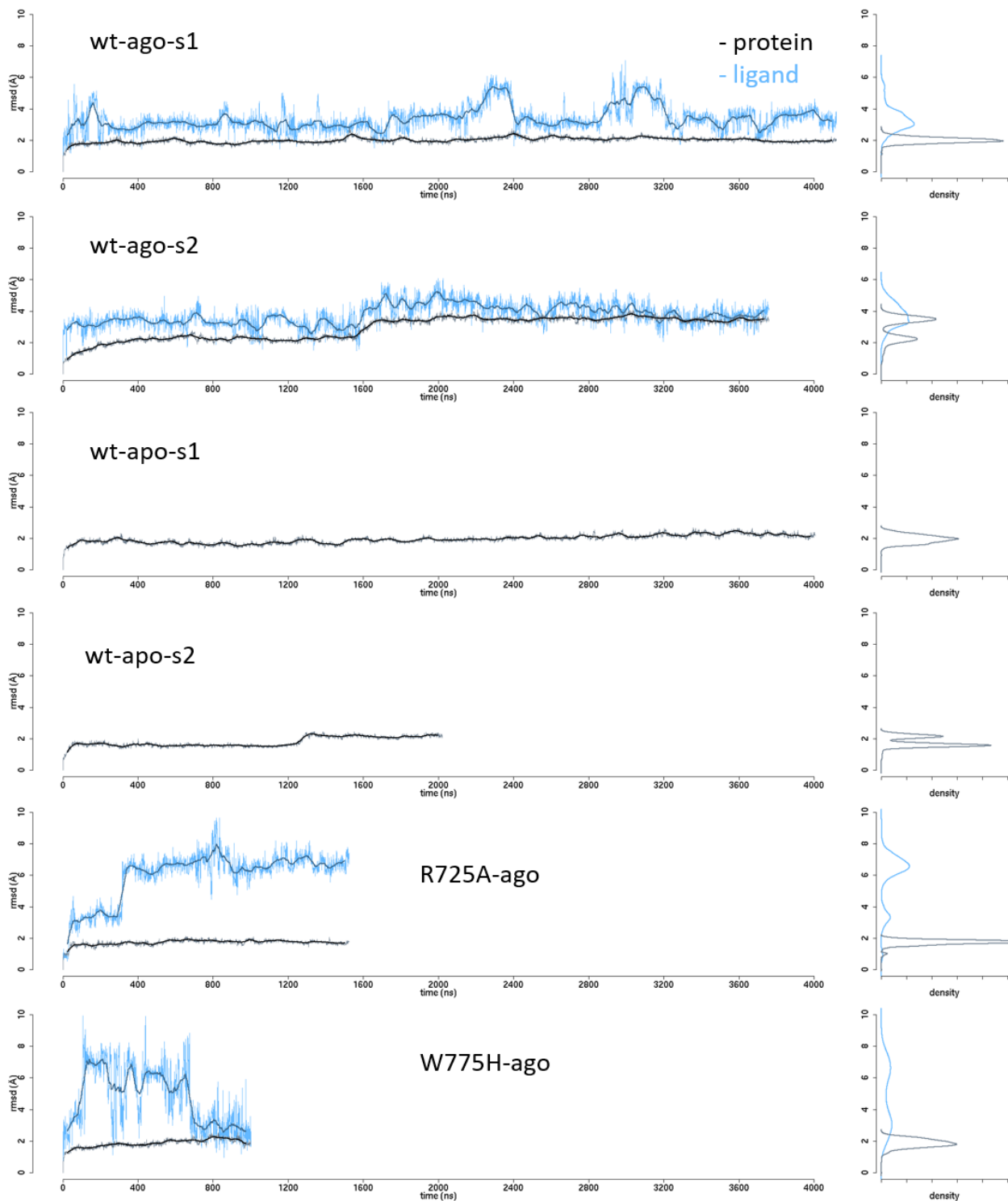
## Figure s2: Trajectory analysis of the molecular dynamics simulations.

Time series plots and density distribution of protein (black) and ligand (blue) RMSD in multiple independent MD runs. Despite the high RMSD values, cyclamate remains within the T1R3 allosteric pocket all along the MD trajectories. It adopts alternate conformations and interacts with different charged or polar residues (Figure s4). See Table s2 for acronym details.

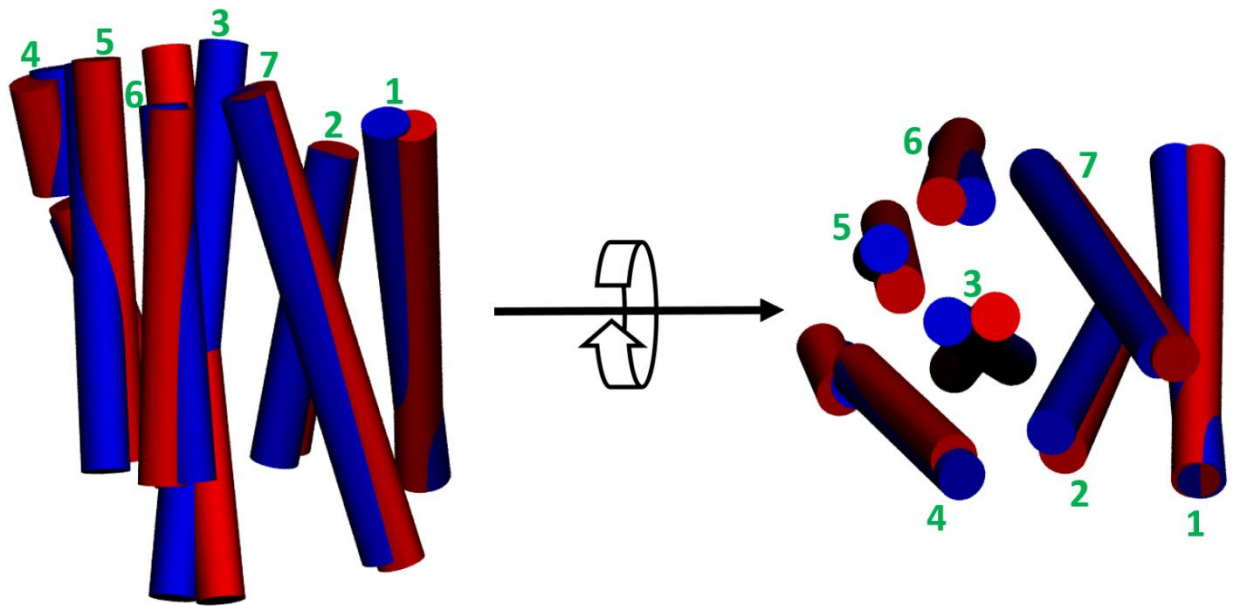Figure s3: Representative structures of wt-ago-s2 MD simulation.

Representative structures of the wild-type receptor bound to cyclamate in the s2 MD simulation. Blue and red structures correspond to the first and the second part of the MD trajectory respectively. Major differences involve TM3, 5 and 6 of the receptor. They may correspond to conformational changes between inactive and active-like states.

**wt-ago-s1**

**wt-ago-s2**

**R725A-ago**

**W775H-ago**

- cyclamate − Q636
- cyclamate − Q637
- cyclamate − R790
- cyclamate − R725
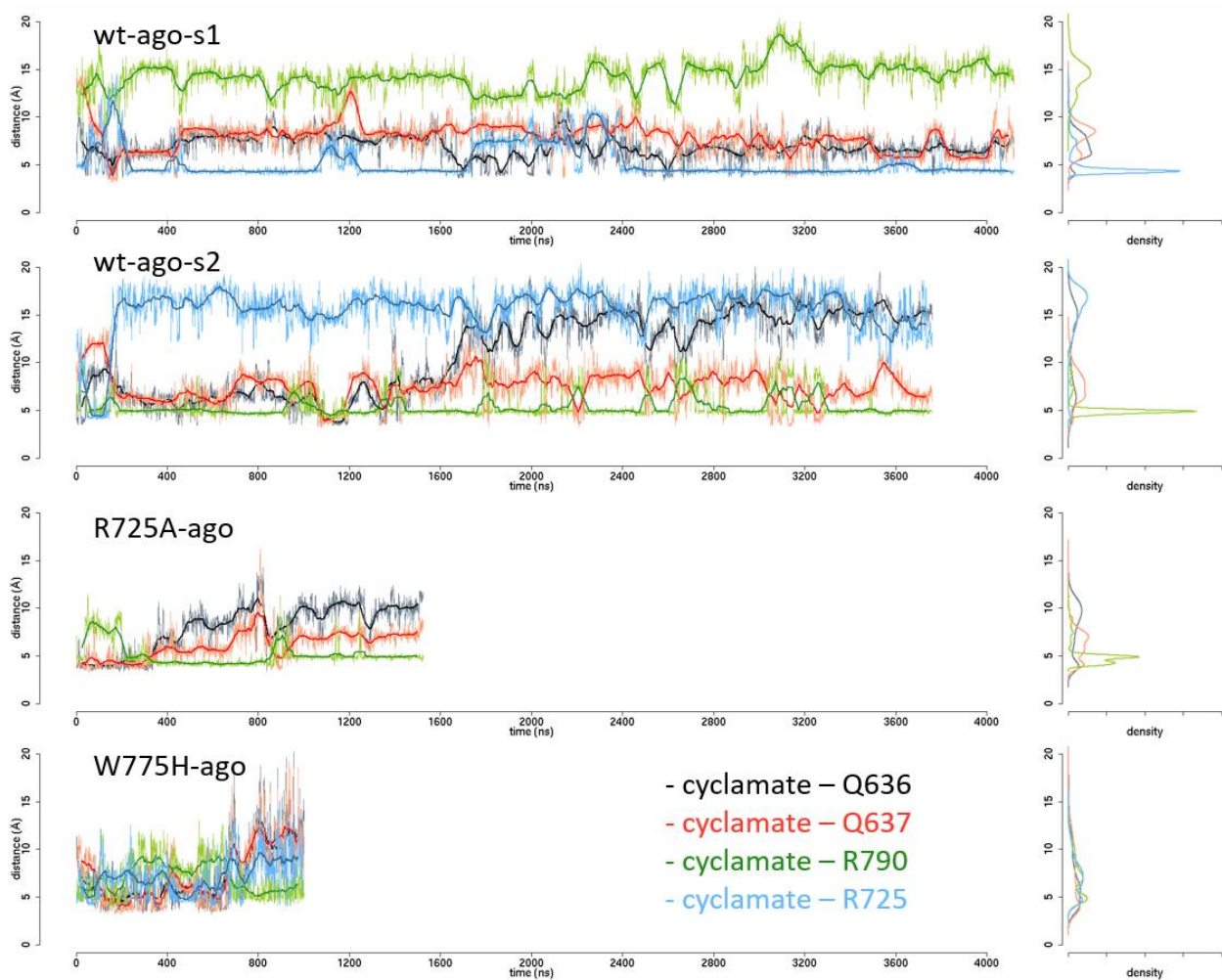
Figure s4: Structural analysis of ligand-receptor interactions.

Time series plots and density distributions of the ligand-receptor distances in *wt-ago* and *mut-ago* MD simulations. To avoid bias in the analysis due to the rotation of the cyclamate sulfamic acid functional group, the sulfur atom has been considered to calculate the ligand-receptor distance. For the same reason, the $C^\zeta$ and $N^{\varepsilon 2}$ atoms have been considered for glutamine (Q636 and Q637) and arginine (R790 and R725) residues respectively. A distance in the range of 4.5 to 5 Å corresponds to a hydrogen bond interaction.

Figure s5: Structural analysis of key residues involved in the T1R3 molecular switch.
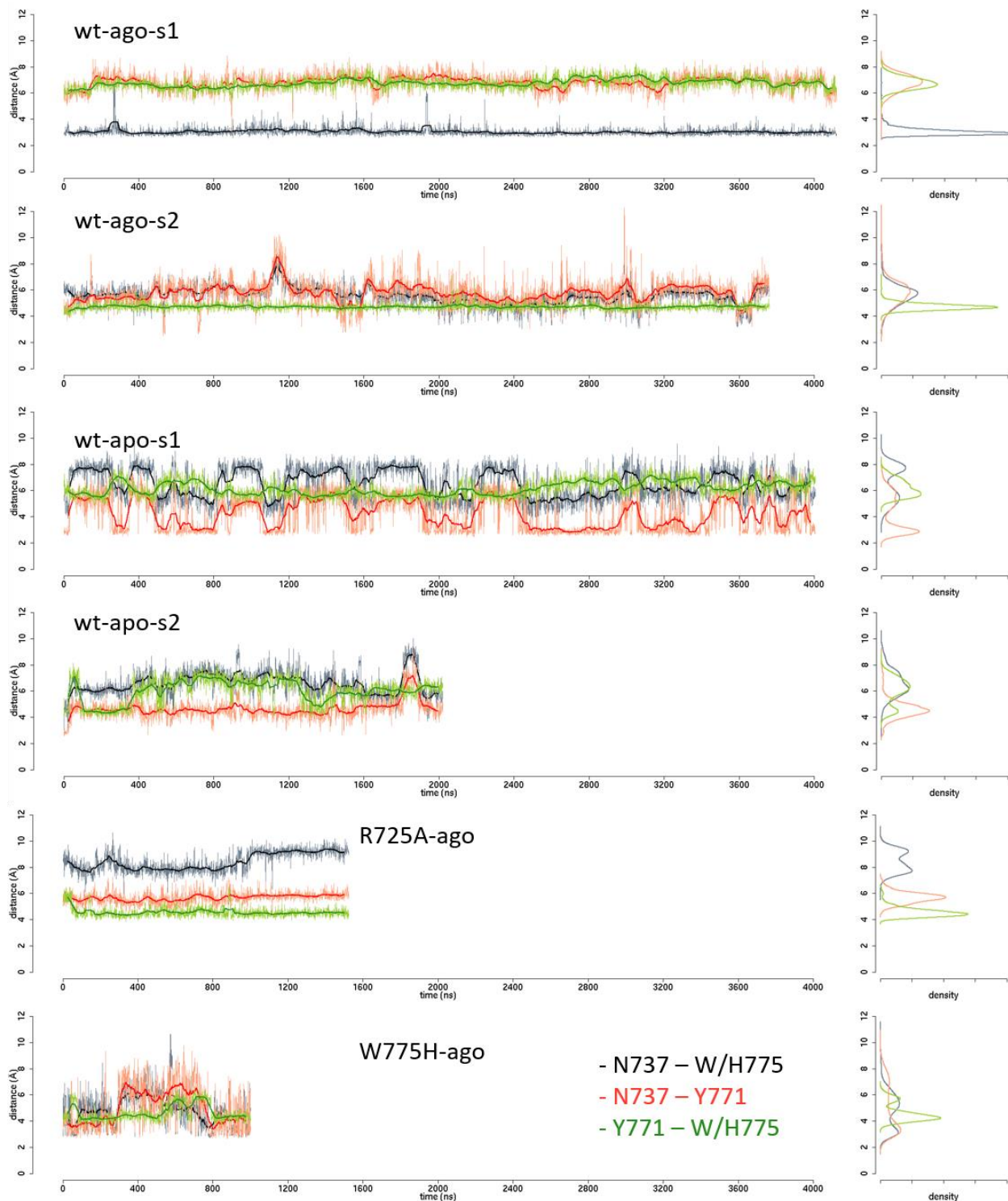
Time series plots and density distributions of N737$^{5.47}$-W/H775$^{6.50}$ and N737$^{5.47}$-Y771$^{6.46}$ distances and Y771$^{6.46}$-W/H775$^{6.50}$ side-chain center-of-mass distance in multiple independent MD simulations (see Table s2 for acronym details). Distances have been calculated between the O$^{δ1}$, O$^{η}$ and N$^{ε1}$ (or N$^{ε2}$ in the case of W775H mutant) atoms for N737$^{5.47}$, Y771$^{6.46}$ and W775$^{6.50}$, respectively. The Y771$^{6.46}$ center of mass has been calculated considering the aromatic ring including the C$^{ε1}$, C$^{ε2}$, C$^{δ1}$, C$^{δ2}$, C$^{γ}$ and C$^{ζ}$ atoms. The W775$^{6.50}$ center of mass has been calculated considering the pyrrole ring including the N$^{ε1}$, C$^{ε2}$, C$^{δ1}$, C$^{δ2}$, C$^{γ}$ atoms.
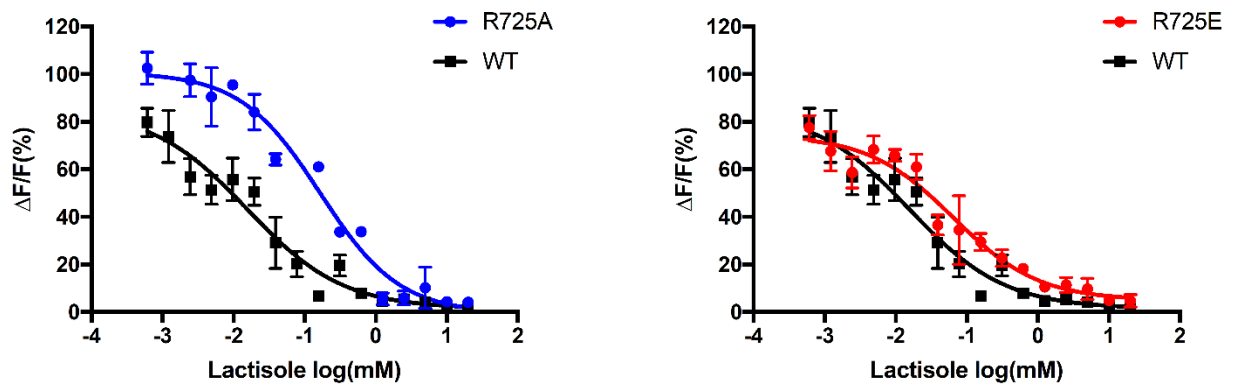
Figure s6: Reduced sensitivity to lactisole of sweet taste receptor R725E and R725A mutants.

Dose-response profiles toward D-tryptophan (10 mM) plus varying amounts of lactisole of (■) wild-type and (●) R725E/A mutant receptors. All transfections were conducted in triplicate.
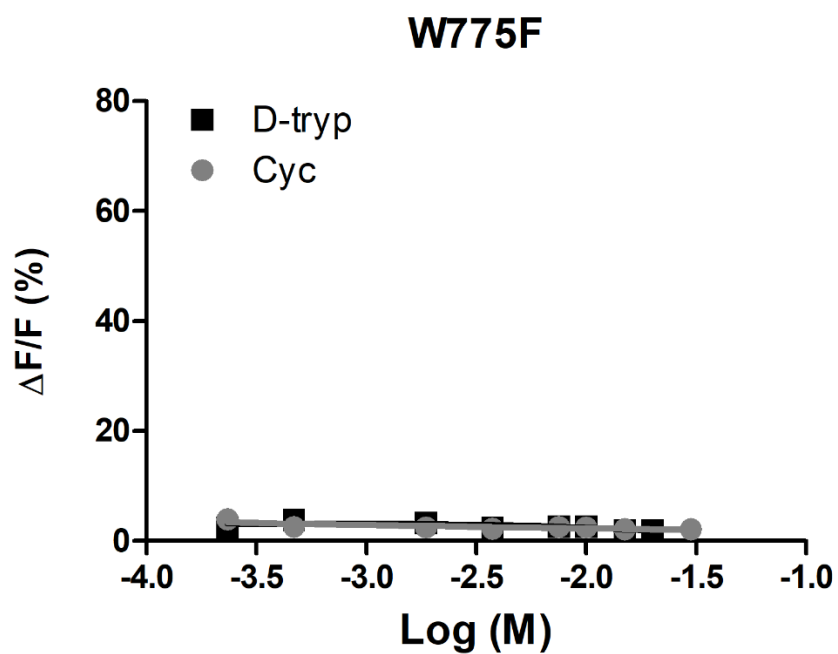
**W775F**

Figure s7: Dose-response curves for sweet taste receptor W775F mutant.

Cyclamate and D-tryptophan dose response curves obtained for the sweet taste receptor single point mutant of residue W775F. All transfections were conducted in triplicate; each experiment was repeated two to three times for D-tryptophan (■) and cyclamate (●).

## Table s3: Meta-analysis of T1R3 TMD site-directed mutagenesis data for allosteric and orthosteric agonists.

Experimental data come from Jiang *et al.* 2005(1) and Winnig *et al.* 2007(2). New site-directed mutagenesis experiments reported in the present study (see Figure 2c) are indicated by an asterisk. Δ(mut-wt) indicates the mutation effects on the receptor response. D-Trp and Asp means D-tryptophan and aspartame respectively.

| Position | Ballesteros Weinstein numbering | Mutation | Δ(mut-wt) | | References |
|---|---|---|---|---|---|
| | | | D-Trp or Asp | Cyclamate | |
| 620 | 2x52 | S⇒A | -36% | -30% | (2) |
| 621 | 2x53 | V⇒I | -36% | -32% | (2) |
| 621 | 2x53 | V⇒L | -94% | -92% | (2) |
| 624 | 2x56 | F⇒L | -79% | -78% | (2) |
| 636 | 3x32 | Q⇒A | -39% | -86% | (2) |
| 636 | 3x32 | Q⇒A | -50% | -104% | (1) |
| 637 | 3x33 | Q⇒E | -48% | -81% | (2) |
| 637 | 3x33 | Q⇒E | -31% | -100% | (1) |
| 640 | 3x36 | S⇒A | -36% | -38% | (2) |
| 640 | 3x36 | S⇒A | -21% | -21% | (1) |
| 641 | 3x37 | H⇒A | -18% | -62% | (2) |
| 641 | 3x37 | H⇒A | -17% | -71% | (1) |
| 644 | 3x40 | L⇒A | -104% | -104% | (1) |
| 645 | 3x41 | T⇒A | -96% | -104% | (1) |
| 699 | 4x50 | Y⇒F | -42% | -49% | (2) |
| 699 | 4x50 | Y⇒L | -30% | -38% | (2) |
| 721 | ecl2 | H⇒A | -39% | -81% | (2) |
| 721 | ecl2 | H⇒A | -38% | -88% | (1) |
| 723 | ecl2 | R⇒A | -30% | -49% | (2) |
| 723 | ecl2 | R⇒A | -17% | -54% | (1) |
| 724 | ecl2 | T⇒L | -30% | -41% | (2) |
| 725 | ecl2 | R⇒A | -12% | -27% | * |
| 725 | ecl2 | R⇒E | -16% | -58% | * |
| 725 | ecl2 | R⇒M | -42% | -35% | (2) |
| 726 | ecl2 | S⇒A | -48% | -76% | (2) |
| 729 | 5x39 | S⇒A | -27% | -38% | (2) |
| 729 | 5x39 | S⇒A | -8% | -29% | (1) |
| 729 | 5x39 | S⇒F | -14% | -15% | * |
| 729 | 5x39 | S⇒T | -41% | -30% | * |
| 729 | 5x39 | S⇒Y | -55% | -70% | * |
| 730 | 5x40 | F⇒A | -21% | -90% | (1) |
| 730 | 5x40 | F⇒C | -63% | -95% | (1) |

| 730 | 5x40 | F⇒L | -30% | -76% | (2) |
|---|---|---|---|---|---|
| 730 | 5x40 | F⇒L | -16% | -82% | (1) |
| 730 | 5x40 | F⇒Q | -32% | -95% | (1) |
| 730 | 5x40 | F⇒S | -37% | -95% | (1) |
| 730 | 5x40 | F⇒T | -47% | -95% | (1) |
| 730 | 5x40 | F⇒W | 0% | -52% | (1) |
| 730 | 5x40 | F⇒Y | 5% | 19% | (1) |
| 733 | 5x43 | A⇒V | -6% | -30% | (2) |
| 733 | 5x43 | A⇒V | 42% | 5% | (1) |
| 734 | 5x44 | H⇒A | -30% | -95% | (2) |
| 734 | 5x44 | H⇒A | -83% | -96% | (1) |
| 735 | 5x45 | A⇒I | 16% | 5% | (1) |
| 737 | 5x47 | N⇒A | -64% | -85% | * |
| 739 | 5x49 | T⇒M | 16% | 18% | (1) |
| 771 | 6x46 | Y⇒A | -104% | -104% | (1) |
| 771 | 6x46 | Y⇒F | -90% | -83% | * |
| 771 | 6x46 | Y⇒L | -19% | -88% | * |
| 775 | 6x50 | W⇒A | -79% | -86% | (2) |
| 775 | 6x50 | W⇒A | -92% | -104% | (1) |
| 775 | 6x50 | W⇒F | -84% | -95% | * |
| 775 | 6x50 | W⇒H | -69% | -90% | * |
| 775 | 6x50 | W⇒Y | -88% | -90% | * |
| 776 | 6x51 | V⇒A | -46% | -38% | (1) |
| 778 | 6x53 | F⇒A | -18% | -70% | (2) |
| 778 | 6x53 | F⇒A | -17% | -88% | (1) |
| 779 | 6x54 | V⇒A | -45% | -30% | (2) |
| 779 | 6x54 | V⇒A | -63% | -54% | (1) |
| 782 | 6x57 | L⇒A | -6% | -81% | (2) |
| 782 | 6x57 | L⇒A | -46% | -100% | (1) |
| 788 | ecl3 | V⇒A | -14% | -18% | (1) |
| 789 | ecl3 | L⇒Y | 50% | 24% | (1) |
| 790 | 7x28 | R⇒A | -72% | -100% | (1) |
| 790 | 7x28 | R⇒E | -52% | -104% | (1) |
| 790 | 7x28 | R⇒H | -100% | -112% | (1) |
| 790 | 7x28 | R⇒K | -28% | -100% | (1) |
| 790 | 7x28 | R⇒Q | -27% | -81% | (2) |
| 790 | 7x28 | R⇒Q | 0% | -94% | (1) |
| 790 | 7x28 | R⇒Y | -60% | -100% | (1) |
| 794 | 7x32 | Q⇒A | -100% | -100% | (1) |
| 798 | 7x36 | L⇒I | -18% | -5% | (2) |
| 798 | 7x36 | L⇒I | 0% | -18% | (1) |
| 800 | 7x38 | L⇒F | -9% | -8% | (2) |
| 800 | 7x38 | L⇒V | 7% | -12% | (1) |
| 801 | 7x39 | C⇒I | -91% | -89% | (2) |
| 801 | 7x39 | C⇒L | -48% | -89% | (2) |
| 802 | 7x40 | V⇒A | -7% | -18% | (1) |
| 804 | 7x42 | G⇒A | -55% | -43% | (2) |
| 804 | 7x42 | G⇒V | -79% | -76% | (2) |
| 805 | 7x43 | I⇒A | -100% | -104% | (1) |
| 807 | 7x45 | A⇒V | 50% | 24% | (1) |
| 808 | 7x46 | A⇒T | 7% | -12% | (1) |

# Figure s8: Multiple sequence alignment of mammalian T1R3 transmembrane domains.

The sequences of 41 mammalian organisms have been obtained and aligned on the UNIPROT webserver. Each full organism name is mentioned at the end of the sequence. The multiple sequence alignment is colored according to percentage identity. The consensus sequence is annotated to display the predicted transmembrane helices (in red), the binding cavity residues (in blue), the transmission switch (in green) and the other molecular switches (in yellow). Red boxes refer to less conserved binding residues as mentioned in Table s1.

# Figure s9: Comparison of allosteric binding pocket of class C GPCR.

T1R3 allosteric binding cavity superimposed on the crystal structure of the mGluR1 (PDB code 4OR2).(Wu et al., 2014) Docked cyclamate and neohesperidine dihydrochalcone (NHDC) are represented by purple and blue sticks respectively while the ligand of mGluR1 (FITM) is represented by green sticks. Residues N737$^{5.47}$, Y771$^{6.46}$ and W775$^{6.50}$ forming the cradle of the allosteric binding site are represented by grey sticks.
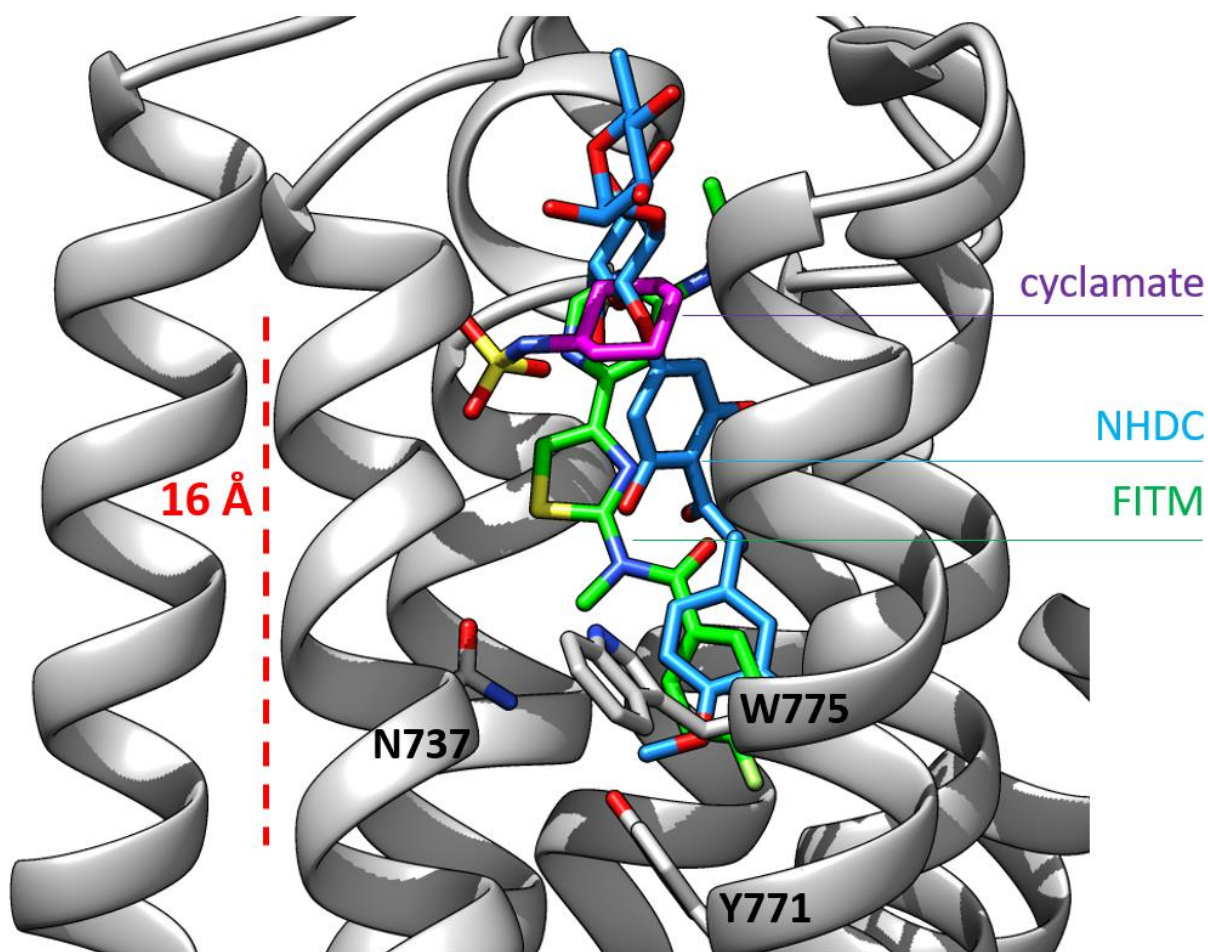
# References

Jiang P, Cui M, Zhao B, Snyder LA, Benard LMJ, Osman R, Max M, and Margolskee RF. 2005. Identification of the cyclamate interaction site within the transmembrane domain of the human sweet taste receptor subunit T1R3. J Biol Chem. 280:34296–34305.

Winnig M, Bufe B, Kratochwil NA, Slack JP, and Meyerhof W. 2007. The binding site for neohesperidin dihydrochalcone at the human sweet taste receptor. BMC Struct Biol. 7:66.

Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, et al. 2014. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. Science. 344:58–64.